SPIM Supplement
Data Quality Objectives

Table of Contents

**This Page Intentionally
Left Blank**

SPIM SUPPLEMENT

DQODATA QUALITY OBJECTIVES

DQO.A   *CERCLIS Data Quality*

In response to issues raised in the last IG Audit, the CERCLIS modernization task was initiated to address improvements to CERCLIS. Part of the CERCLIS modernization effort focuses on the development of a comprehensive data quality program to ensure we have accurate, timely, and accessible information to track, assess, and plan the Superfund program. One essential component of this data quality program is the development of Data Quality Objectives (DQOs) for Superfund information. DQOs define the tools and procedures used to ensure that the data entered into WasteLAN and the data reported from CERCLIS is complete, accurate, timely and consistent.

DQO.B   *What are Data Quality Objectives?*

DQOs identify and document information needed for decision-making, as well as requirements for data quality, specifically how complete, accurate, timely and consistent the data needs to be to support the decision-making process. The Information Management Branch (IMB) has identified an approach for this effort based on the Agency's DQO guidance. This approach involves drafting DQOs for priority areas of CERCLIS first (i.e., GPRA is the top priority); meeting with appropriate data sponsors to review the draft documents; and finalizing the documents by incorporating insights gleaned during the meetings.

The purpose of each DQO is to:
- define the decision/accomplishment and who is responsible for the decision,
- explain the procedures for recording the decision in the database,
- identify the scope (the extent or range of sites considered),
- identify all of the relevant data elements required in order to receive credit,
- define the specific data criteria and/or valid values for each data element required in order to receive credit,
- describe the data quality requirements for timeliness, accuracy, completeness and consistency.

Specifically, the Data Quality Objectives define how CERCLIS/WasteLAN data will be evaluated based on objective assessment of each of the following measures:

- the extent to which the data covering a specific interval of Superfund program activity are promptly entered into CERCLIS/WasteLAN. (Current/Timeliness)
- the extent to which the required data are reported and recorded in the database. (Complete)
- the extent to which the values of the data elements use the standard definitions or codes and the extent to which these definitions and codes are used in the same way by all users. (Consistent)
- the extent to which the data recorded in CERCLIS reflect the correct, true, or reported values as reflected in hard copy documents. (Accurate)

The goal is to define data quality objectives that can measured and reviewed using audit reports. Many of the data quality objectives are defined in the Superfund Program Implementation Manual (SPIM). Recently, a "timeliness" chart was added in Appendices A and B to define the required documentation for specific targets and measures as well as the timeliness for entering data from the source document into WasteLAN. The DQOs will be used to document the specific data quality objectives for each Superfund accomplishment. If there is conflict between a SPIM definition and a DQO the SPIM defination will prevail. Hence the phrase at the bottom of each page of this supplement: "This Document Does Not Supercede the SPIM"

DQO.C   Defining the Data Quality Objective Measures

Current:

>Current refers to the status of information in the database – as measured by the length of time between the actual event (or receipt of documentation about the event) and its official entry in the database. CERCLIS/WasteLAN standards for data to be considered current vary by the type of data being entered into the system. Most actions that are used for accomplishment reporting should be entered into CERCLIS within 5 working days of the receipt of the information, even though the GPRA measure is reported annually. Some program areas do not require such stringent data entry requirements. In some circumstances, data should be entered within 10 working days of the receipt of the information, or within a month, or even quarterly prior to the quarterly data pull.

Complete:

>Completeness refers to the amount of required data present in the database at a specific point in time. Complete data assures that all pertinent information is available for use when needed. The Superfund Program Implementation Manual (SPIM) and the SPIM Coding Guide provides the minimum set of Nationally Essential Data Elements (NEDE) required for CERCLIS/WasteLAN.

>Current and complete are closely related and are often considered together. In particular, completeness must always be evaluated with respect to currency – that is, data are complete (or incomplete) as of today. In some instances, however, it is easy to separate these two factors. For example, a site assessment report may be received and the data entered into CERCLIS/WasteLAN in a timely manner but the entered data lack required elements. In this case, information would be considered current; but not complete.

>WasteLAN includes prompts to assist users in entering complete information, and incorporates procedures that ensure that accomplishment credit is given only if all the required information exists. There are also QA reports that pull and display sites/actions for which the necessary information is not entered.

Consistent:

>Consistency refers to the extent to which appropriate values are used for a data element as defined nationally. For management reports to be the most effective, data must be comparable over time within the area of interest. Comparable codes or values must be used for the area of interest. They must be used for the same data elements over time and in different geographical area if valid comparisons are to be made by Superfund managers. For example, program managers may want to identify facilities with active pump and treat technology.

>Consistency errors usually originate with the program or facility staff who prepare source documents. They differ from accuracy errors in that the value is correct as far as the coder is concerned, but the selection of the code is not consistent with the approved definition of the data element, the data element value, or with the way the code is used by others.

Accurate:

>Accuracy refers to the absence of erroneous data resulting from mistakes during any point in the data preparation, entry or transmission process. Errors sometimes result from mistakes by key-entry personnel, but program or facility personnel who prepare the source documents used for data entry can also introduce errors. Data entry errors are usually misspellings and incorrectly entered values while transmission errors often result in transposition of characters and dropped digits. All data entered into CERCLIS/WasteLAN for reporting targets and accomplishments must be supported by objective evidence, that is, documentation that a third party could examine and arrive at the same conclusion (objective evidence rule). Dates and data in the source document should match the data that is entered into WasteLAN.